Original Article

IJCIH
Converging Healthcare & Technology

INTERNATIONAL JOURNAL OF CONVERGENCE IN HEALTHCARE

Published by
IJCIH & Pratyaksh Medical Care LLP
www.ijcih.com

Novel Architecture for Reducing Sports Injuries in Football

Durgansh Sharma[1], Vaibhava Sharma[2]

[1]Associate Professor,University of Petroleum & Energy Studies (UPES), Department of Cybernetics & School of Computer Science, Energy Acres Building, Bidholi, Dehradun–248007, Uttarakhand, India, [2]MBA-ESG (Sports Management), Jain University (Deemed to be University), Bengaluru, Karnataka–560009, India

Abstract

This paper proposes a novel architecture to reduce sports injury in the game of football. The research framework non-invasively intrudes the captured postures of athletes during their respective playing sessions using computer vision and wearable devices. The inclusion of computer vision transforms the postures and gestures as semantic moves, which after processed using the proposed framework predicts the injury in advance depending upon the moves of the respective player.

Keywords: *Sports Injury, Football, Computer Vision, Gesture Recognition, Predictive analytics.*

Introduction

As per the healthcare data regarding physical injuries, it found that more than twenty percent of accidental injuries getting treatment at various medical facilities in the world are associated with sports and wrong postures during related training sessions.[6] It dilutes the career progression of various sports personalities passionate to groom themselves in the field of their favourite sports football. To reduce these injuries requires implementation of solutions towards the prevention of injuries. The sports injury prevention requires the incorporation of sports and exercise specific intrusions of structural measures towards education of sports persons, and then its safety promotion includes the predecessor and extended drives that are required towards success of the required measures. Comprehensive sports safety promotion programmes thus require that the perspective on the sports injury problem becomes broader than consideration of the individual athlete.

Review of Literature: The videos of the football game are most important component towards this research work and understanding them comprehensively is always a challenging task, as the game comprises of action sequences of game with time called temporal action localization. The detection of players of both teams, referee and linemen is again a challenging task, which needs object detection through object tracking. As we know most of the work towards understanding a video mainly, focus on some isolated aspects towards the video analysis and setup a correlation among the task post-identification and localization.[1]

Next step is towards benchmarking of human activity as objects in video or players in the game of football where

Corresponding Author:
Dr. Durgansh Sharma
Associate Professor, University of Petroleum & Energy Studies (UPES), Department of Cybernetics & School of Computer Science, Energy Acres Building, Bidholi, Dehradun–248007, Uttarakhand, India
e-mail: durgansh.sharma@ddn.upes.ac.in
https://www.upes.ac.in

ActivityNet has supported by covering a wide range of players activities that are of interest the game of football. ActivityNet provides samples from 203 classes of human activity bearing an average of 137 untrimmed videos per class and 1.41 activity instances per video, for 849 video hours. This really helped in identifying various postures and gestures that could be threatening towards a sports injury.[2]

Next step is to identify the kinetics of human actions where re-evaluation is always required to establish the validation of that action towards injury prone moves. Identifying using a video is again challenging task where Kinetics has enough data, with approximate 400 human action classes and over 400 clips per class, and is collected from realistic, challenging videos for various actions including sports. Than analysis, using the architectures towards the task of action classification on the dataset check the performance improvement towards benchmarking datasets after pre-training on Kinetics. A two-Stream Inflated 3D ConvNet (I3D) that is based on 2D ConvNet inflation that uses different filters and pooling kernels of deep image classification ConvNets which are expanded into 3D, made possible to provide seamless spatio-temporal feature extractors from video while leveraging successful ImageNet architecture designs and even their parameters. [3]Next requirement for this paper needed another important dataset for spotting action in different frames of football games played and recorded. The repository of "soccernet" comprises of 500 complete football games from six main European leagues, covering three seasons from 2014 to 2017 with duration of 764 hours. The important part is 6,637 temporal annotations automatically parsed from online football match reports at a one-minute resolution for three main classes of events (Goal, Yellow/Red Card, and Substitution). [4]As we know, the number of cameras are increasing day by day in the real-time events like football game. Furthermore, the requirement of background subtraction for football video sequences was satisfactorily taken care by "ViBE" a universal background subtractor.[5]

This research work is towards predicting sports injury in the game of football for the same medical perspective is most important towards the analysis. For the same evidence from Medical Journal of Australia, Victoria were taken into account towards the sports injury in various games including football.[6] The analysis and prediction towards dynamic sports like football a research work towards time-to-event analysis was satisfactorily provided in the research work from British Journal of Sports Medicine.[7]

Methodology

The proposed model captures the moves of football player while they are practising and even during the real game happening. It based on the modality that the poseof a player staged initially to fill the energy in muscles, for opening the style or move to unload through signature gesture orposture for the expected outcome. Then every part of the event recorded to enhance the footballplayer's outcome or predict for an injury beforehand. It includes the cognizance of wearable sensors implanted in their outfits, shoes etc. and the posture recognition techniques using computer vision. The proposed model is able to capture relevant information from the gestures or postures of the sports person relevant to respective performance; it will be a key to success for the sports personalities and reduce the issues for many sports events. The proposed model comprises of integrated applications for providing relevant information towards their training sessions. It collects their respective moves and providespre-processed feedback towards injury prevention. It will be helpful towards decision making for the player selection in a team, as it is required for enhancing the strategic deployment against the competitors.

During the training session, it extracts the information from the video frames captured by high precision cameras. Initially, the computer vision algorithm designed to identify individual football players from both teams using facial points and gait recognition. Then, it maps players move with the alignment Region of Interest. Furthermore, Semantic Segmentation with the help of conv-layers of Deep Convolution Neural Network (VGG-19), Instance Segmentation segments different instances of players, such as labelling players from each teamplaying with different colours. The frameworks also trained to identify the players, goalkeepers, linemen and umpire.

The process subdivided into different phases towards implementation of VGG-19; every phase has to be fine-tuned to provide the desired results. As we know that, the concept of deep learning depends upon the usage of convolutional neural networks. The Phase one (Figure 1) and two (Figure 2) are reflecting the process of convolution layer extracting features from the given frame or matrix of image of a football game. Individual identification and feature extraction during the scheduled time of game provides the needful information.

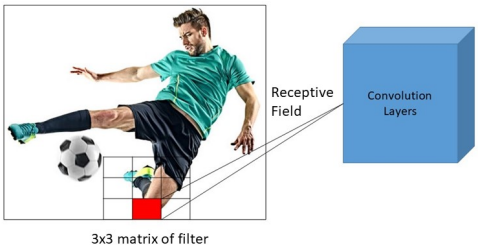

Figure 1: Phase One of Convolution Layer

When the cameraman provide the real-time image, the algorithm applies over the same and process of convolution and pooling initiates as mentioned in the flow chart (Figure 5).

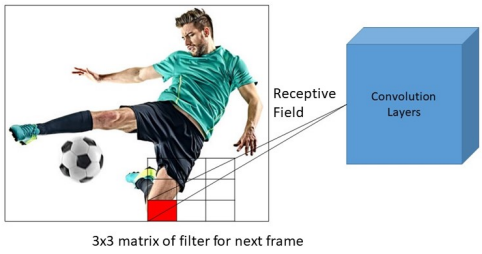

Figure 2: Phase Two of Convolution Layer

The process of convolution comprises of dot product of image matrix and the filter matrix to get the output matrix with computed data in every cell as mentioned in Figure 3 below.

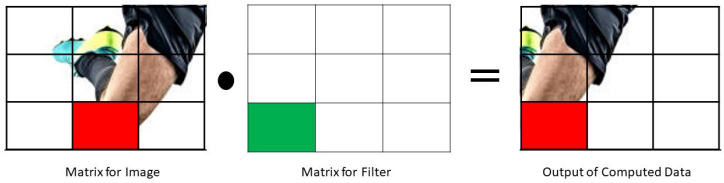

Figure 3: Process of Convolution Computation

Finally, the process of pooling of cells is been taken care by the pooling layer (Figure 4) of VGG-19 framework used for this procedure.

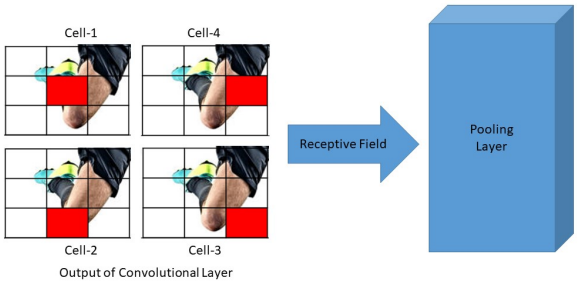

Figure 4: Process of pooling the outcomes of Convolutional layer

The flowchart (Figure 5) describes the convolution procedure used for getting computed values and pooling procedure that comprises of pooling data then the process of maximum value retrieval and finally computed data. The proposed architecture (Figure 6) comprises of input from training data with correct gestures and postures of individual player for playing the game of football stored as database. The feature extraction process got a feedback network from deep convolution neural network (VGG19), these are real-time feature extraction which is on continuation basis.

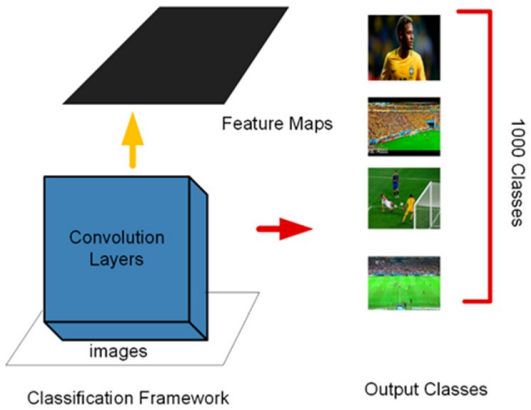

Figure 5: Flow Chart towards feature extraction of football players

The external data and the query image uploaded by the camera operator shall further update the database with tags provided by auto-tagging system comprised of correct or faulty posture of the player.

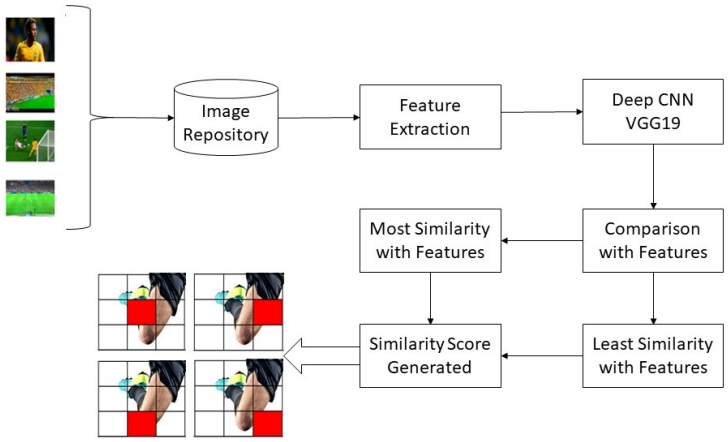

Figure 6: Architecture of Proposed Framework for Reducing Sports Injuries in Football

The extracted feature data after comparison finally identified with the most similar images and least similar images not considered for the further processing of the pose of the player (Figure 7). During classification, there is generally an image with a single object as the focus of the recommender system and the related task is to identify the image. In order to segment instances, as needed to carry out issues that are more complex. We see complicated sights of the game with multiple overlapping players as objects and different backgrounds, and the system not only classify these different objects but also identify their boundaries using edge detection, differences, and relations to one another during the game.

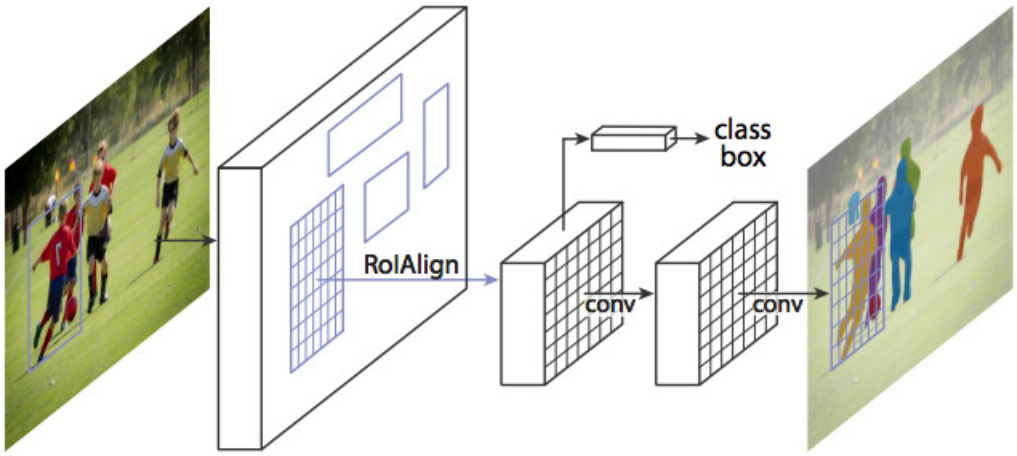

Figure 7: Scenario Implementation of proposed architecture over clip from football game

When the regions of the feature map of players wither overlapped or non-aligned from the respective original image. The Deep CNN solves the issue by adjusting region of interest (RoI) for providing efficient outcome of players.

Results

The dataset used from SoccerDB [1], a novel large-scale corpus of manual annotations for football.

Figure 8: Wrong posture detected in stick framework

The dataset constitutes the most inclusive dataset for football video, with more than 3 million annotations, different computer vision tasks and multiple benchmark results. It is helpful in performing the research to fetch results. The proposed architecture including four stages were distinctly taken into consideration for collecting the action of the player, concentrating on the posture while playing, the drive towards releasing the energy for any action of the football game, and finally recording all the needful actions and integrate them to analyse and predict. The action of the player mentioned in figure 8 is showing the predicted feedback against the posture taken by player, the result where damage or injury predicted is encircled.

Ethical Clearance: Taken

Source of Funding: Self

Conflict of Interest: Nil

References

1. Jiang Y, Cui K, Chen L, Wang C, Xu C. SoccerDB: A Large-Scale Database for Comprehensive Video Understanding. In: MMSports'20. 2020. p. 1–8.
2. Heilbron FC, Escorcia V, Ghanem B, Niebles JC. Activitynet: A large-scale video benchmark for human activity understanding. In: IEEE Conference on Computer Vision and Pattern Recognition. IEEE; 2015. p. 961–70.
3. Carreira J, Zisserman A. Quo vadis, action recognition? a new model and the kinetics dataset. In: IEEE Conference on Computer Vision and Pattern Recognition. IEEE; 2017. p. 6299–308.
4. Giancola S, Amine M, Dghaily T, Ghanem B. Soccernet: A scalable dataset for action spotting in soccer videos. In: IEEE Conference on Computer Vision and Pattern Recognition Workshops. IEEE; 2018. p. 1711–21.

5. Barnich O, Van Droogenbroeck M. ViBe: A universal background subtraction algorithm for video sequences. IEEE Transactions on Image Processing,. 2011;20(6):1709–24.
6. Tharanga Fernando D, Berecki-Gisolf J, Finch CF. Sports injuries in Victoria, 2012-13 to 2014-15: evidence from emergency department records. The Medical Journal of Australia [Internet]. 2018;208(6):255–60. Available from: https://www.mja.com.au/journal/2018/208/6/sports-injuries-victoria-2012-13-2014-15-evidence-emergency-department-records
7. Nielsen RO, Bertelsen ML, Ramskov D, Møller M, Hulme A, Theisen D, et al. Time-to-event analysis for sports injury research part 2: time-varying outcomes. British Journal of Sports Medicine. 2019;53(1):70–8.